Survey of Digital Tools Used By Each Group to Automate Systematic Reviews (including Rapid, Overviews, Living etc)

20 November 2025

By Alex Sutton

(with “help” from LLM Copilot to compile results ☺ then ☹)

Response Rate

- I got 15 completed forms from groups
- I don’t know what the denominator is(!) but this seemed a good response rate to me
  - Thank you all who contributed to this!

Automation Tools Summary: Scoping

Task	Core Tool(s)	Sometimes Used	Exploring
Scoping	Google (+Scholar) x6, Copilot x2, Undermind (ABo) Elicit (ABo) Scite (ABo) PROSPERO, Elicit, OSKR EPPI Reviewer with judicious use of its LLM connections (OpenAI models, Deepseek, Llama, Mistral)	LLMs ChatGPT x5 Microsoft Copilot and/or Claude (for scoping out search terms) Gemini overview (Google) Elicit AI x 2 Undermind x2 ResearchRabbit x1 Covidence Rayyan Connected Papers Litmaps Consensus We separate tool from technology, so sometimes look at LLMs using APIs and code things up quickly in python.	Tools in TERA x4, Next gen LMM tools including ALLEN.AI x 3 Undermind x2 Elicit x2 Wordfreq Connectedpapers Ai assisted topic mapping / semantic clustering Nested Knowledge Citation Chaser Research Rabbit Connected Papers PubReMiner • Polyglot Search Translator • Yale MeSH Analyzer LLMs such as ChatGPT Asta/Find Papers Using embeddings to help cluster records and visualise relationships

Automation Tools Summary: Protocol Writing

Task	Core Tool(s)	Sometimes Used	Exploring
Protocol Writing	(LLM for plain language abstracts) x2 Google Notebook LM Claud (PA)	LLMs ChatGPT x2, Copilot to check grammar and language use Occasional use of generative AI (Copilot, to improve wording, or suggest plain language alternatives) Edinburgh LLM for summaries/abstracts Claude AI JBI SUMARI Cochrane Review Manager (Overleaf)	LLMs x3

Automation Tools Summary: Searching

Task	Core Tool(s)	Sometimes Used	Exploring
Searching	Citationchaser x3 Connected papers Google Advanced TERA: SpiderCite Polyglot De-duplication conventional databases + OpenAlex (via OpenAlex tools in EPPI Reviewer) Automation tools in EPPI Reviewer to maintain living evidence syntheses with OpenAlex as the main source.	Undermind x2 Elicit x2 DownthemAll for website capture Google Scholar, CORE, Citationchaser, Scite, EPPI-Reviewer (OpenAlex) NotebookLM (search strategy development) Experiment with 'agentic' tools such as Perplexity, Asta, (Elicit)	CiteSource API-based large-scale search federated search linked NIC / NHS libraries TERA Nested Knowledge Citation Chaser Research Rabbit Connected Papers PubReMiner Polyglot Search Translator Yale MeSH Analyzer Interested in deduplication programme Asta/Find Papers Impact of paywall on new tools such as Elicit

Automation Tools Summary: Screening Abstracts

Task	Core Tool(s)	Sometimes Used	Exploring
Screening Abstracts	Rayyan x6 Covidence x4 EndNote x3 EPPI-Reviewer x4 (prioritising records for screening, classification (e.g. RCT Classifier, custom/ bespoke ML classifiers)) Laser AI Covidence Nested Knowledge – AI Robot reviewer (use the AI robot reviewer as second reviewer for the remaining 80%). Human reviewer resolves the disagreements between the robot and the person. (Massively time saving)	Rayyan x6 Covidence x5 EPPI Reviewer x3 (including priority screening tool & use of classifier models) (Using LLMs) EndNote JBI SUMARI Covidence Eppi-Reviewer Covidence Nested Knowledge x 2 (Using probability of advancement before you turn the robot reviewer on (i.e. during the duplicate screen) (performance depends on simplicity of review question – if complex still useful for obvious excludes) LLMs	EppiReviewer x5 (inc. GenAI) & (especially in the context of mapping reviews) https://asreview.nl/ (priority ordering) x2 Covidence x2 EasySLR x2 DistillerSR Abstrackr LLM (We are interested in how LLMs could be used to increase efficiency without compromising integrity) Rayyan Nested Knowledge ActiveSLR

Automation Tools Summary: Screening Full Text

Task	Core Tool(s)	Sometimes Used	Exploring
Screening Full Text	Covidence x4 EPPI-Reviewer x3 EndNote x2 Rayyan x4 (for recording decisions, usually in the Ti&Ab screening tab) Laser AI	Covidence x5 Rayyan x3 EndNote JBI SUMARI Eppi-Reviewer Nested Knowledge Covidence LLMs Evaluating using the various LLMs in EPPI Reviewer	EPPI Reviewer x5 (interested in how LLMs could be used to increase efficiency) & (especially in the context of mapping reviews) Covidence x2 DistillerSR ChatPDF Humata LLM EasySLR Nested Knowledge ActiveSLR

Task	Core Tool(s)	Sometimes Used	Exploring
Data Extraction	Covidence x3 EPPI-Reviewer x2 MS Forms x2 MS Excel x2 Laser AI MS Access WebPlotDigitizer Google translate (for translation from other languages) Qualtrics Maxqda Zotero	ChatGPT x2 (to support human data extraction by identifying relevant information within academic papers and other wider sources) Eppi-Reviewer x2 (eval. LLMS in it) JBI SUMARI Elicit LLMs Copilot (tested but find that both ChatGPT and Copilot can be inaccurate in extracting certain details) Claude AI Google Notebook LM Google Gemini Adobe PDF's AI tools (to support human data extraction by identifying relevant information within each paper) Nested Knowledge (tested AI-based data extraction in but chose not to continue at this time) Covidence Rayyan	LLMs x4, Chat GPT 5.1 research analysis f function Interested in how LLMs could be used to increase efficiency, including query construction for focused data extraction (retrieval-augmented generation) Rayyan x 3 EPPI Reviewer x2 Humata ChatPDF EasySLR Covidence Nested Knowledge EasySLR ActiveSLR Involved in a large project that is developing a 'data extraction evaluation toolkit': a platform to build and evaluate the use of language models + agents for data extraction / RoB tasks across domain

Automation Tools Summary: RoB/Quality assessment of studies

Task	Core Tool(s)	Sometimes Used	Exploring
RoB/Quality assessment of studies	JBI tools Critical Assessment Tool (CAT) MS Excel Macro (RoB 2 tool) Covidence	Eppi-Reviewer x 2 (Evaluating using the various LLMs in EPPI Reviewer) Drummond checklist JBI SUMARI MS Excel (ROBINS-E tool) Cochrane Review Manager Claude AI Google Gemini	LLMs x5, (to increase efficiency) Web-based RoB 2 tool

Automation Tools Summary: Data Synthesis – Quantitative

Task	Core Tool(s)	Sometimes Used	Exploring
Data Synthesis – Quantitative	STATA x6 R/RStudio x6 RevMan x3 (plots only) GitHub (version control) x2, Python SAS EPPI Reviewer	CRSU Apps x2 R x2 (for presenting overlap in umbrella reviews) Online meta-analysis tools LLMs ChatGPT or Copilot (may be used to help check narratives) ChatGPT for R troubleshooting Meta-analysis – JBI SUMARI WinBUGS STATA EPPI mapper (for visual presentation of evidence maps)	CRSU Apps x 2 R with AI augmentation using Retrieval-augmented generation (RAG) R for meta-analyses ActiveSLR (GN)

Automation Tools Summary: Data Synthesis – Qualitative

Task	Core Tool(s)	Sometimes Used	Exploring
Data Synthesis – Qualitative	Microsoft Office Word (used for grouping codes and themes) Zotero Maxqda OSKR NVivo Mostly NA x 2 EPPI Reviewer	NVivo x3 (considered the AI features available for NVivo 15 (AI-assisted auto coding) but found they were too limited for our needs within a large-scale document analysis.) Claude AI x2 Copilot to help group codes	LLMs x3 (planning retrospective comparison with existing synthesis) & (looked at using for some synthesis / scoping tasks) R with AI augmentation using Retrieval-augmented generation (RAG)

Automation Tools Summary: Certainty of Evidence Assessment

Task	Core Tool(s)	Sometimes Used	Exploring
Certainty of Evidence Assessment	GRADEpro x4 CINeMA x3 (NMA) CERQUAL x3 (Qual) CAT Mostly NA EPPI Reviewer	iSOQ - GRADE-CERqual	GRADEpro x3 (AI being added) GRADE-CERQual, Interested in how LLMs could be used

Automation Tools Summary: Reference Managing

Task	Core Tool(s)	Sometimes Used	Exploring
Reference Managing	EndNote x8 Zotero x4 Rayyan x1 Mendeley x1 Endnote Citationfinder EPPI Reviewer	INSPECT-AI (checks reference lists for retractions) TERA for deduplicationn EndNote Cochrane Review Manager Zotero x2 (for citation management and tools) & (use of Zotero tools in EPPI Reviewer to semi-automatically retrieve full-text reports)	Ai Augmented Ref management Zotero

Automation Tools Summary: Report Writing

Task	Core Tool(s)	Sometimes Used	Exploring
Report Writing	Canva CoPilot (PES) Hemingway app Doctools Manually using Microsoft Office Word templates	LLMs (ChatGPT or Copilot to help with language) (ChatGPT for PPIE summary) (generating or revising the plain language summary) (improve wording, or suggest plain language alternatives) Cochrane Review Manager	LLMs x2 Gemini/custom gems

Automation Tools Summary: Dissemination

Task	Core Tool(s)	Sometimes Used	Exploring
Dissemination	Specific websites x7 HCRW EC website PHW website Website (https://bristol-esg.org/) BlueSky x3 Twitter/X x3 LinkedIn x2 Medrxiv Social media Notebook LM Blogs YouTube TikTok Instagram Facebook CANVA	LLM x3 (ChatGPT for revising plain language / lay audience summaries, policy briefs) (for generating or revising the plain language summary) Canva for posters Prezi for infographics Padlet for AI image generation AI generated audio transcript Claude AI	X BlueSky LinkedIn AI assisted lay summary Data visualisation platforms Infogram creation. Exploring: https://scispace.com/ https://www.napkin.ai/ x1, Google Notebook LM

Additional Tasks Added To Table

Task	Core Tool(s)	Sometimes Used	Exploring
Integrity checking for RCTs	None	Starting to evaluate AI version of INSPECT-SR (called INSPECT-AI)	Further evaluation of this tool with other evidence synthesis groups.

Task	Core Tool(s)	Sometimes Used	Exploring
Cost-effectiveness modelling	None	None	Gemini (DP x1, mostly for working out function errors)

What tools would you like training in?

- General LLM
- Automation tools in EPPI Reviewer
- EPPI reviewer; It would be great to be able to use LLM to create blogs/summaries of our publications etc but currently no access to paid versions together with uncertainty over security of the information uploaded
- How to use the tools and how to evaluate them using living SWARs and other methods research studies.
- Also general training on how to use automation tools appropriately and how to validate them for use in a specific context.
- R + (with AI)
- Use of LLMs when including evidence in different languages, beyond only English language.
- Quality or risk assessment without compromising integrity/validity
- INSPECT-SR and INSPECT-AI protocol
- Tools for AI assisted qualitative coding
- AI tools to assist with development of search strategies and data extraction

What tools would your RSG possibly be able to provide training in?

- Automation tools in EPPI Reviewer (!) – How to use the tools and how to evaluate them using living SWARs and other methods research studies.
- Also general training on how to use automation tools appropriately and how to validate them for use in a specific context.
- Support for using Rayyan for screening and reference management.
- Laser AI (but requires a license/subscription, and developers also provide training)
- AS Review
- Nested Knowledge, but it is a paid service and we don't have access to the full AI functionality as that would be very expensive!

Any tools you have had bad experience with (i.e. not fit for purpose)?

- **AI:**

- ChatGPT and Copilot can hallucinate and be inaccurate for data extraction.

- ChatGPT for R troubleshooting/coding solutions, mostly for data processing and results visualisation, keeps suggesting arguments that are not recognized by the function/package being used

- Covidence, I personally did not find it useful. However, I have not used this recently, so I am unsure if they have improved their AI functions.

- Google Gemini - Inconsistent data extraction from the same articles.
- CoPilot/Claude - Producing infographics
- Chat GPT in basic form

- Fundamentally, very few automation tools currently provide much in the way of an evidence base to support their use. (E.g. Elicit)

Can you recommend any good resources/training on any of the tools you use?

- The training the Rayyan team provided was quite helpful if accessing the paid version, they also have some guides available online.
- There are also good Cochrane resources and also useful webinars on Evidence Synthesis Ireland
- Review Manager online training
- AS Review has work great for me. I believe that by being an open source is also quite accessible and there are different open resources that can be used. https://asreview.nl/
- Resources for INSPECT-SR here: https://osf.io/b74wj/overview
- The EPPI Centre provides numerous self-directed training resources for EPPI Reviewer 6: https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3822
- EPPI Centre staff also provide synchronous training opportunities and are available to answer further questions.
- NK Documentation
- Covidence online training
- Edinburgh Uni training on ELM use was very helpful
  (https://information-services.ed.ac.uk/computing/comms-and-collab/elm)

Helpful Links Supplied

- As well as share some of the expertise I had while co-supervising a PhD student who researched on methods to automate systematic reviews. Here some of her publications if this is of interest

https://www.sciencedirect.com/science/article/pii/S0933365724002318

https://link.springer.com/chapter/10.1007/978-3-031-66538-7_26

https://link.springer.com/article/10.1007/s10462-024-10844-w

https://aclanthology.org/2024.naacl-long.206/

- And other work done in this area:

https://www.sciencedirect.com/science/article/pii/S2772662223000024

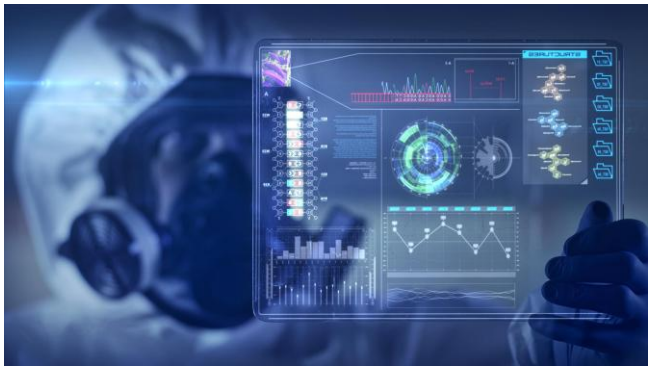
https://onlinelibrary.wiley.com/doi/full/10.1002/jrsm.1553

https://www.valueinhealthjournal.com/article/S1098-3015(22)01409-7/fulltext

Key Insights

- Most Common Tools by Category:
  - Screening: Rayyan, Covidence, EasySLR, EPPI Reviewer
  - Searching: Google Scholar, PubMed, Polyglot Search Translator
  - Data Extraction: Covidence, EasySLR, Nested Knowledge
  - Reference Management: EndNote, Zotero
- Emerging Trends:
  - Rapid adoption of LLMs (ChatGPT, Claude, Gemini) for scoping and protocol writing
  - Increased interest in AI-assisted screening and semantic clustering tools
- Training Needs:
  - High demand for training in LLMs and AI tools (ChatGPT, Elicit, Undermind)
  - Requests for guidance on EPPI Reviewer and advanced screening platforms

Tools respondents would like training in



AI-Powered Language Models

Training interest in LLMs like ChatGPT and Microsoft Copilot supports tasks such as scoping and data extraction in reviews.

Screening and Data Management Tools

Respondents seek training in EPPI Reviewer, Rayyan, and Covidence for efficient screening and data handling.

AI Literature Discovery Tools

Specialized AI tools such as Elicit and Undermind aid literature discovery and semantic clustering, requiring user guidance.

Statistical and Synthesis Software

Training demand includes statistical tools like RStudio and STATA for quantitative data analysis in evidence synthesis.

Tools RSGs could provide training in

Screening and Data Management Tools

RSGs offer structured training on platforms like EPPI Reviewer, enhancing skills in review mapping and classifiers.

Onboarding and Troubleshooting Support

Expertise in Rayyan and Covidence allows RSGs to assist users with onboarding and resolving technical issues.

AI-Assisted Platforms and Reference Management

Training includes AI-assisted screening tools like Nested Knowledge and reference managers such as EndNote and Zotero.

Foundational vs Specialized Training

While foundational tools are covered by RSGs, advanced AI-driven technologies may require external specialized training.

Tools with negative experiences

Inaccurate AI Outputs

AI tools like ChatGPT sometimes produce hallucinated or unreliable data during extraction and coding tasks, causing trust issues.

Limitations of Screening Platforms

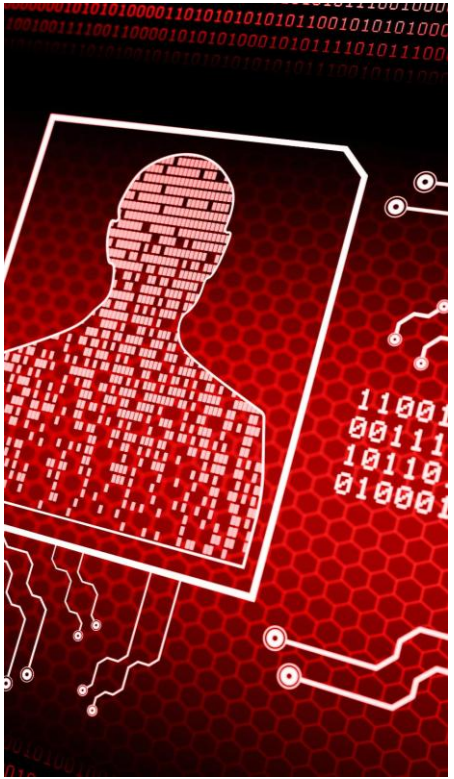
Users faced challenges with Covidence due to limited AI features and usability problems for complex systematic reviews.

Inconsistent Data Extraction

Google Gemini showed inconsistencies in extracting data from identical sources, impacting review accuracy.

Suboptimal Performance in Specialized Tasks

Tools like Claude and CoPilot underperform in generating structured outputs despite potential in summaries and infographics.



Recommended resources and training materials

Comprehensive Digital Tool Tutorials

Covidence and Review Manager provide detailed online tutorials and documentation for mastering core functionalities.

EPPI Centre Training

The EPPI Centre offers self-directed modules and live sessions focused on EPPI Reviewer users' needs.

AI-Assisted Screening Platforms

ASReview is an open-source platform with accessible guides and community resources for AI-assisted screening.

Academic and Documentation Resources

University-led programs and published research provide practical insights into AI summarization and automation methods.